Air Force Studies & Analyses Agency

MASTR
A New Look

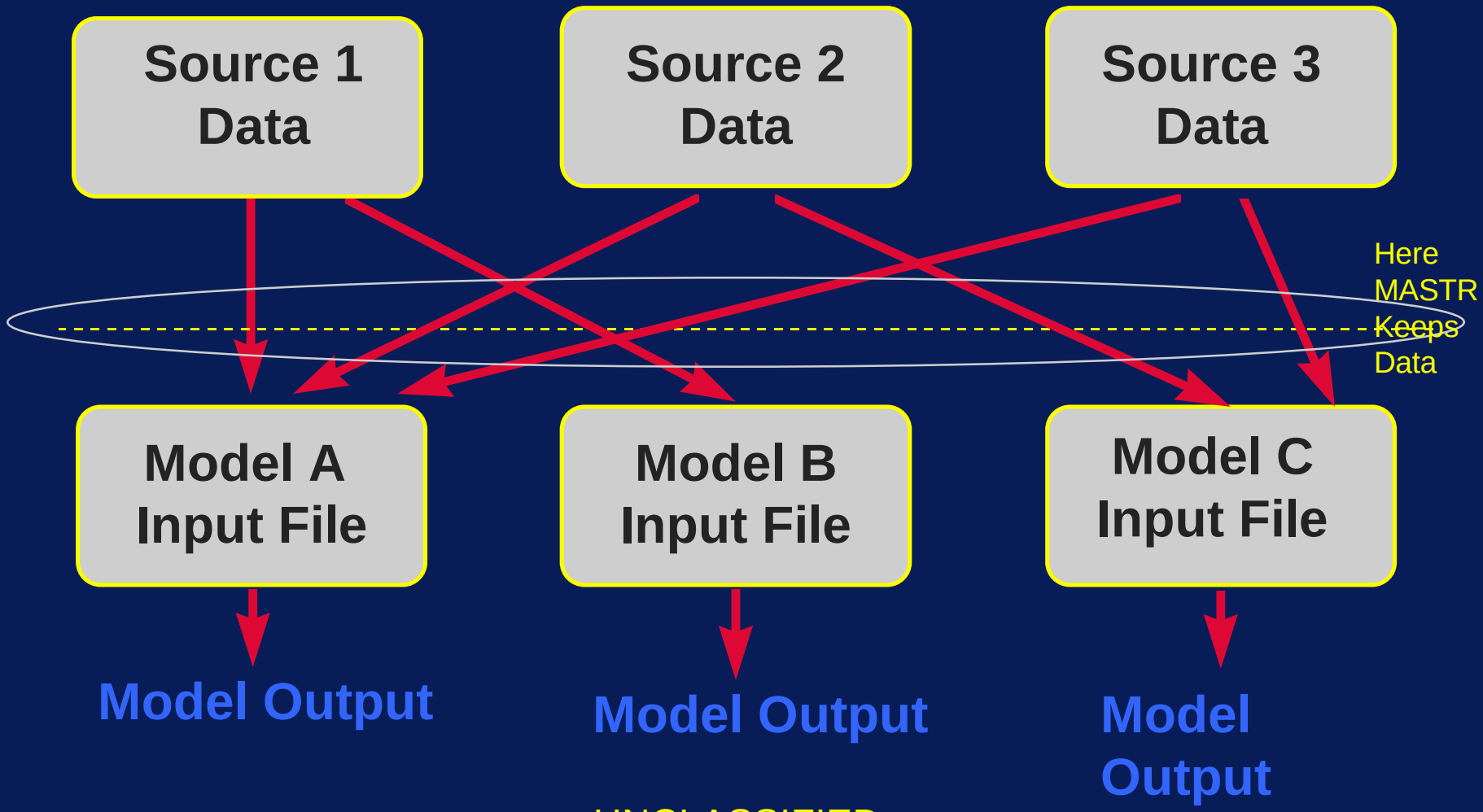
Database Redesign

Burvin Jenkins
CACI

Database Facilities For MASTR

- Access to consistent data is the primary reason for MASTR's existence
- MASTR must access data from many data sources
- Current MASTR implementation uses a relational database (Oracle)
- Object Oriented style used in current data design

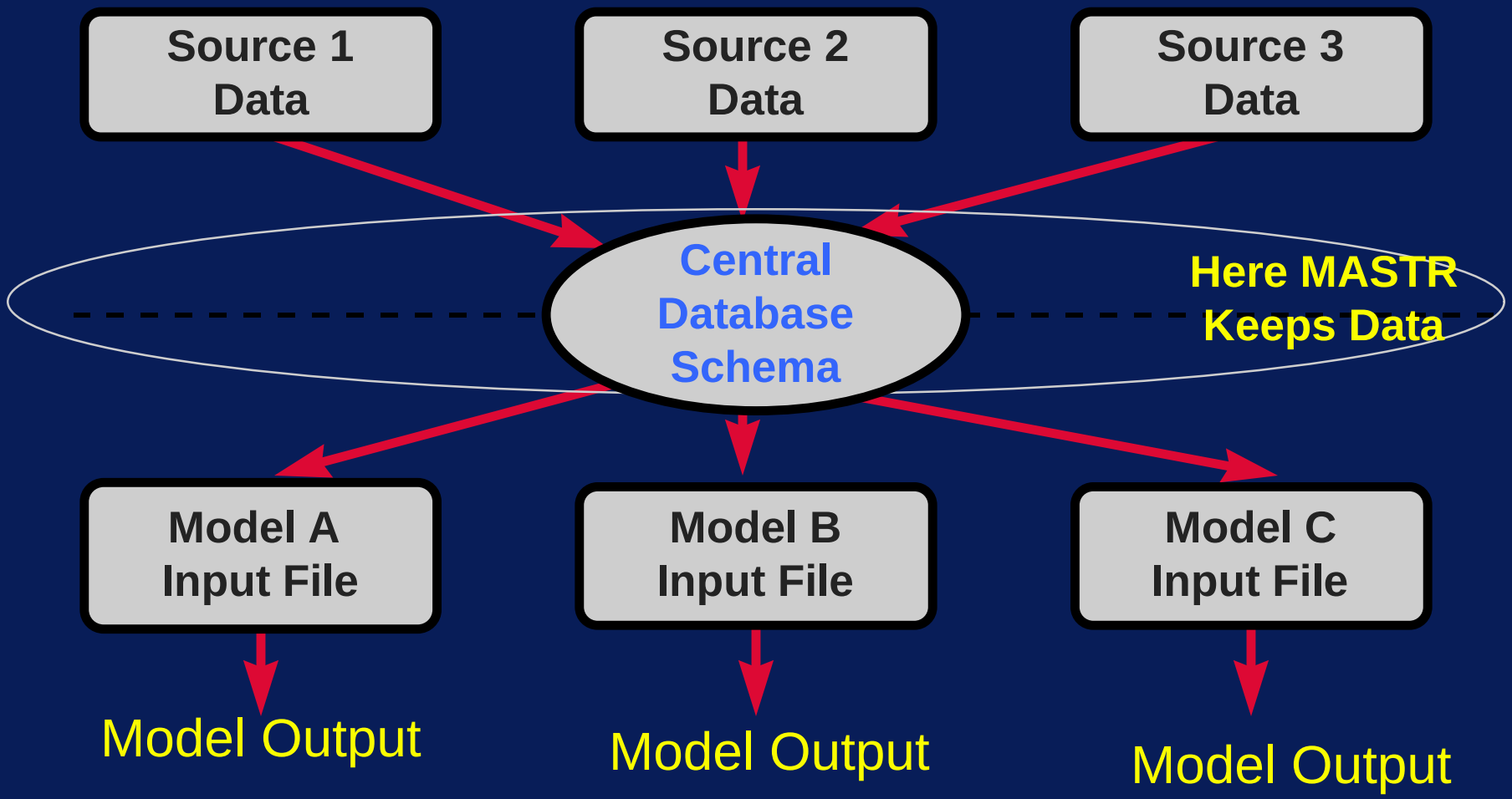
Model Support



Current - Single Schema Approach

- "Central Database" is the structure used in the current version of MASTR
- A single schema is used to define
  - the data structure and set of relationships
  - for all data sources and models
- Data from each source must be converted to fit the structure of the Central Database schema

Current Data Style



Problems with Single Schema

- Long delays in integrating new data sources
- Poorly structured tables (e.g. tables with many NULL columns)
- Different interpretations of a field depending on the data source of the record

Reasons for problems

- Each data source specifies a data structure and set of relationships which may be *unique to that source*
- There is no such thing as a *typical* data source
- There is no *universal frame of reference, set of relationships, or structure* that can be used to describe all data sources

New - Multiple Schema Approach

- Replace “Central Database” with “Central Repository”
- Divide the database into separate schemas, each optimized for a particular data source or model
- Encapsulate the necessary relationships between schemas into “views”

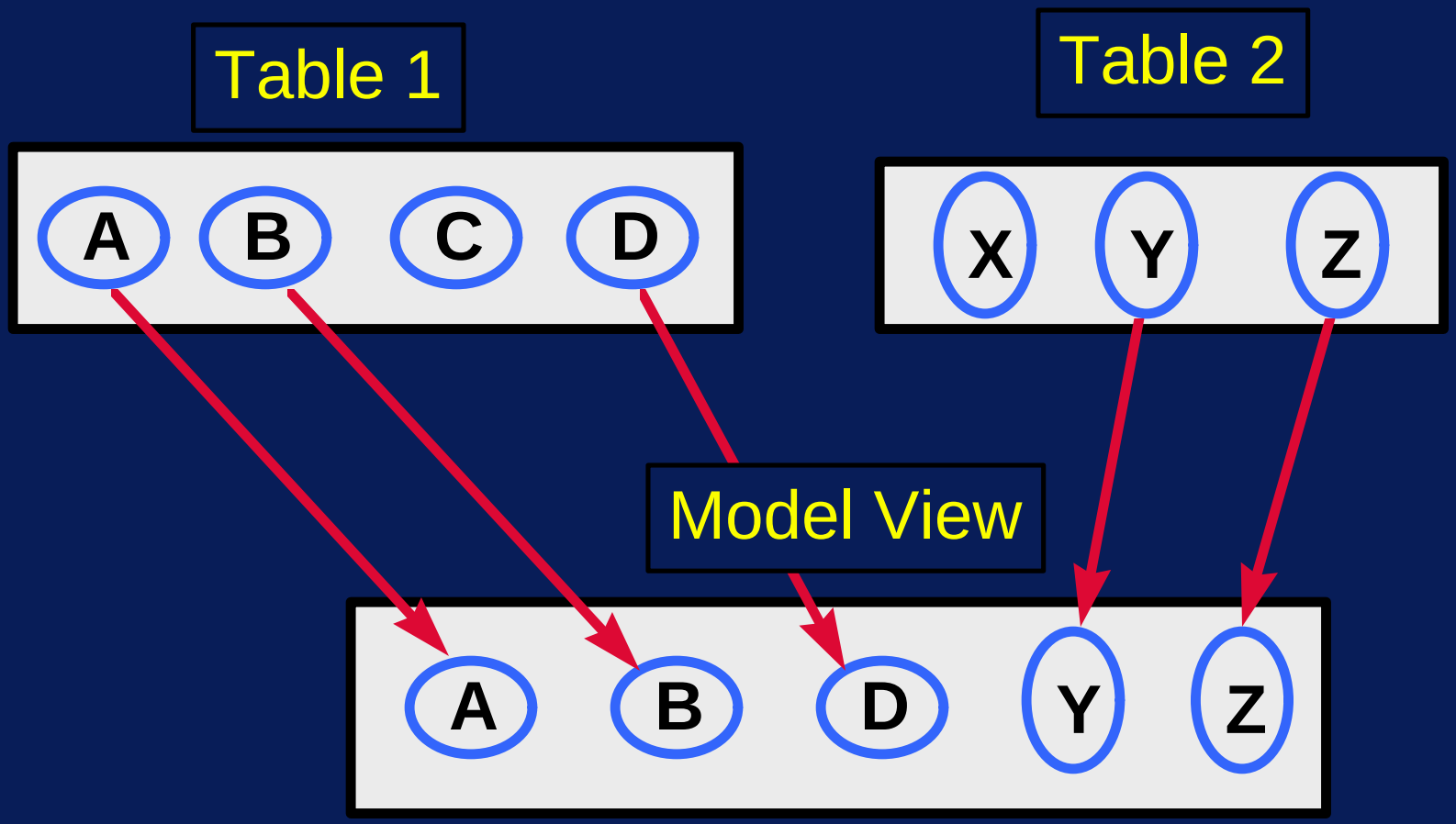
How New Approach Works (Data Side)

- For each data source, a schema is constructed
- Tables created within the schema define the relationships expressed within the data source and its associated documentation
- Each schema is designed so we can make best use of the database by how we store and index the data

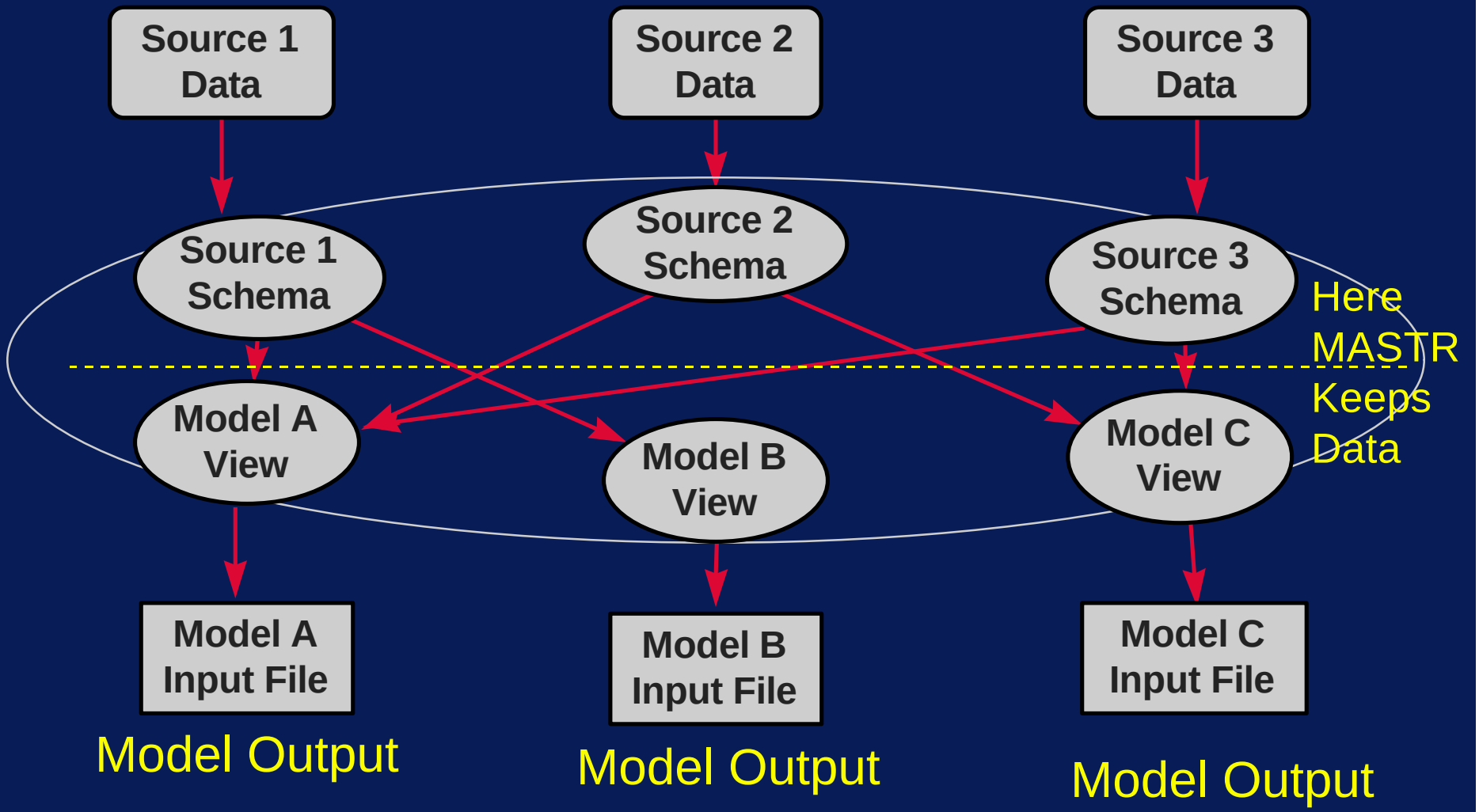
How New Approach Works (Model Side)

- For each model, a schema is constructed
- Model schema should consist of “views” which map to the corresponding data source tables
- A “view” is a stored query that provides a custom-tailored presentation of the data in one or more tables

Creating A Model View



New Data Style



Reasons for using views

- Views provide a layer of abstraction between the data source and the query that references the view
- Each record within the view corresponds to a combination of fields from other tables in the database
- Views allow relationships to be defined between independent schemas

Problems still to be overcome!

- Loading a data source requires significant interpretation of the data (e.g. comment fields that contain critical pieces of information)
- Some fields must be split to normalize relationships (e.g. unit subordination code)
- Link tables will be created to relate information between different data sources (e.g. MSFD air units<--->IDB air bases)

Where are we now ?

- Created a data model for the MSFD - IDEF1X compliant
- Using ERwin data modeling tool (purchased Oracle Designer 2000)
- Tool allows for creation of ER diagram and generates ODBC SQL
- Need to translate ODBC SQL into Oracle SQL and port code to Unix platform

Conclusion

- The single schema concept, although simple, is neither flexible nor efficient
- The multiple schema concept reduces the data complexity problems and allows for easier integration of “new” data sources
- The model support provided by the database never changes!